**🛍️ Store Sales Forecasting Project**

This project aims to forecast daily sales for a chain of stores using historical data, holidays, oil prices, and transaction logs.

**📁 Dataset Overview**

The project uses multiple CSV files:

* train.csv: Historical daily sales data for training.
* test.csv: Test dataset with dates and stores for which sales need to be predicted.
* stores.csv: Metadata about each store (type, location, etc.).
* oil.csv: Daily oil prices.
* holidays\_events.csv: Includes national and local holidays, events, etc.
* transactions.csv: Number of daily transactions per store.

**🧠 Goal**

To predict the **unit sales** for each store on each date in the test set.

**🛠️ Project Steps**

1. **Data Loading**  
   All CSVs are loaded using pandas.
2. **Data Cleaning**
   * Interpolation is used to fill missing values in the oil price data.
   * Date columns are converted to datetime format for consistency.
3. **Feature Engineering** *(to be added if present)*
   * Extracted features like day, month, year, or holiday flags.
4. **Model Building** *(to be added if present)*
   * You can add details about regression models, XGBoost, or any other approach used.
5. **Evaluation** *(to be added if present)*
   * RMSE, MAE, or other metrics used to evaluate model performance.

**📦 Requirements**

* Python 3.x
* Pandas
* NumPy
* Matplotlib / Seaborn (for EDA, if used)
* scikit-learn or other ML libraries

You can install the requirements via:

bash

CopyEdit

pip install -r requirements.txt

**📊 Future Improvements**

* Hyperparameter tuning
* Time-series models (like ARIMA or Prophet)
* Incorporate weather data